DOCUMENT RESUME

ED 349 708 EA 024 328

AUTHOR Rieck, William A.
TITLE Helping Teachers Improve Student Assessment Procedures: The Supervisor's Role.
PUB DATE Apr 92
NOTE 11p.; Paper presented at the Annual Meeting of the Association for Supervision and Curriculum Development (New Orleans, LA, April 4-7, 1992).
PUB TYPE Guides - Non-Classroom Use (055) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Administrator Role; Elementary Secondary Education; *Evaluation Methods; Informal Assessment; *Student Evaluation; Teacher Administrator Relationship; Teacher Supervision; *Test Construction; *Testing

ABSTRACT

Guidelines for the supervisor's role in improving teachers' assessments of their students are offered in this paper. Suggestions for improving formal assessment are provided first, with attention to the various types of formal tests, general test design, and grading procedures. Ways to informally monitor students' learning processes are briefly described next. Suggestions for the administrator's role in supervising teachers' student evaluation processes are offered in the final section. (Contains 24 references.) (LMI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

HELPING TEACHERS IMPROVE STUDENT ASSESSMENT PROCEDURES: THE SUPERVISOR'S ROLE

William A. Rieck, Ed.D.
Head, Department of Curriculum and Instruction
University of Southwestern Louisiana

ED349708

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

W Rieck

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EA 024 325

BEST COPY AVAILABLE

HELPING TEACHERS IMPROVE STUDENT ASSESSMENT PROCEDURES: THE SUPERVISOR'S ROLE

William A. Rieck, Ed.D.
Head, Department of Curriculum and Instruction
University of Southwestern Louisiana

Introduction

Assessment is very much on the minds of administrators, teachers, parents, and students. National concerns have been expressed over student performance and the way we measure the performance of students. With this is mind it is very appropriate to consider the state of assessment and how the administrative and supervisory members of ASCD can help teachers develop fair methods of evaluating student performance.

In the strict sense of the terms, **assessment** refers to determining what students have learned while **evaluation** refers to attaching a quality symbol to the results of assessment. For practical purposes we will use the terms synonymously. The concepts and suggestions on assessment found in this booklet are intended to be reminders of some of the material covered in our workshop. This booklet is not the definitive word on the issue and for this reason a selected bibliography is included at the end.

Importance

The psychological impact of grades is significant, specially at the elementary levels. Children who bring home poor grades are frequently chastised by their parents. Such negative feedback from schools and parents may lead to a lessening of self-esteem and confidence which could result in a downward spiral of performance. If we must assign grades, therefor, it becomes very important to make certain their meanings are understood and that their assignment is fair and realistic.

Grades are also determinators. The grades student receive are used to determine promotion, course qualification, and even college admission. They are so important that Stiggins (1988) reported that teacher directed assessment was the most important issue in assessment today. Stiggins estimated that from 20% to 30% of a teacher's time is spent on assessment related activity; such time commitment reflects the importance and the difficulty of student evaluation.

There has been some research that shows grades, though used as determinators and predictors, are not necessarily valid for that purpose. Andrews (1984) followed the GPA trend and the SAT/ACT trend among incoming college freshman from 1969 through 1979 and found that the GPA of the incoming students consistently progressed, while the SAT/ACT continually declined. In a more recent study by Moralles (1992) it was shown that grades

ED 021 328

were not good predictors of results in quiz bowl competitions while ACT scores were excellent predictors. These studies tend to suggest that either there is something missing with respect to curriculum, or that grading is not realistic.

Reasons for inadequate assessment may be many. It has been shown that most teachers never have a course in tests and measurements (Schwartz and Lisstz, 1987; Popham and Hamberton, 1990). In a recent study of all NCATE accredited institutions (Rieck, 1991) is was shown that 57% of the institutions did not require course work in tests and measurements and that over half admit their students have less than complete preparation in formative and informal assessment procedures.

It is , perhaps, not an exaggeration to say that there may be literally hundreds of thousands of teachers in the United States who are relatively incompetent when it comes to assessment. If this is so, it becomes the responsibility of administrators and supervisors to assist those teachers in developing their skills.

Teachers' Needs

There are a few simple principles of which teachers need to be aware. I tell my students these are "Rieck's Rules on Assessment." The principles are:

1) There needs to be both informal and formal formative assessment throughout each unit of instruction.

2) There needs to be at least two different types of formal assessment on each topic or unit.

3) Assessment models must provide for a wide variety of measurements with no single measurement weighing excessively.

4) Students have a right to know exactly how their performance will be evaluated and how their grade will be assigned.

Teachers need to remember that **formal** assessment involves grades while **informal** does not. Homework, for example, is not formal. Homework is the time when students are supposed to make mistakes and to grade such work is counter-productive.

Informal assessment, in addition to homework, can be in the form of guided practice conducted during class, quality teacher questioning, monitoring of body language, and general class discussion.

Formal assessment can include tests (paper-pencil, experiential, and oral); quizzes, projects, reports, presentations, debates, lab experiments, dioramas, class participation, etc. The greater the variety of assessment procedures, the greater the likelihood of realistic evaluation.

Too frequently teachers relay excessively on a unit test as the primary source of data. Generally, if any given evaluative component is valued at 33% or more of a student's grade the component is weighted too heavily. Administrators need to look at the assessment models of teachers to make certain they are realistic.

Students have a real need to know what is expected of them and what it will take to earn a particular grade. Teachers should let students know what expectations are during the first or second day of class.

Tests: A Major Form of Formal Assessment

Tests may be **criterion referenced** or **normative**. **Normative** measures are those where students are compared to other, large groups of students and are not generally appropriate for classroom testing. **Criterion** tests are those where there is a correspondence between test item and objective. Criterion referenced items are most appropriate for classroom testing in a course. It is appropriate to discuss the several types of test items which teachers may use, followed by a discussion of general test design and grading procedures.

True False items are popular with many teachers, but should be avoided. Too often such items are open to interpretation and misjudgment. For example, consider the True/ False statement:

"The Sun rises every morning."

When I ask my students that question they say "true." Of course, the answer is "false" because the Sun never rises, the Earth rotates. After the groans someone will say it was a trick question. Not true. It is scientifically correct, but an elementary teacher may expect a "true" answer based on a non scientific objective!

If true / false items are to be used it is important to employ one of two devices to correct for the guessing factor:

(1) Make the items correctable

(2) Use a correction formula

Correctable true / false simply means the student is expected to indicate why the statement is false or must correct the statement to make it true. This approach reduces the probability of credit by guessing since no credit is awarded for the "false" response unless it is explained or corrected.

A correction formula is a satirically valid means of correcting for guessing. A simple method is to take the number correct and subtract the

number incorrect. If a student leaves a blank, admitting he or she does not know, it is not scored as incorrect. A student who thinks he knows 85 answers out of 100 and then gets 5 incorrect would have a score of "80" if he or she left the blanks, but a score of 60 if he or she guessed incorrectly at the 15 the student was not sure of.

Completion items are valid for low level objectives such as definitions, number facts, and the like. Care must be taken to assure that not too many of these items are used because they can not generally test the higher order thinking skills.

Multiple choice can be excellent items on a test. A good multiple choice item can easily test upper levels of comprehension and, if well thought out, higher levels of cognition as well. There are two important rules in writing multiple choice items:

1) The distractors used should all seem reasonable and appropriate; not frivolous or obviously incorrect.

2) The distractors should be placed in random order, alphabetical order, or numeric order to avoid a "favorite answer" syndrome.

Lets suppose we frame a knowledge level question in history using 4 choices (4 is the minimum acceptable number of choices).:

The first President of the United States to live in the White House was

(A) Mondale
(B) Adams
(C) Jefferson
(D) Nixon

Choices A and D are obviously not acceptable as distractors because Mondale was never a president and everyone knows the White House was in use before the Nixon administration. It would be better to list the first four Presidents in alphabetical order which would produce reasonable distractors in an appropriate order. To place the Presidents in order of their terms would be too much of a hint!

It should be noted that each choice started on a different line. This is for readability. We want to assess what a student knows or understands and do not want to confuse the person with jamned up choice which could lead to overlooking a possibility, albeit you will use more paper. The SAT, ACT and other tests use separate lines, and for good reason. Teacher tests should also use separate lines.

Matching items can be viewed as a type of multiple choice. It is always advisable to have more choices then there are items, to indicate that some

choices may be used more than once, and some choices not used at all. Such procedures will create many choices for each item.

Essay items are good for higher order thinking skills and for testing multiple objectives or inter-relationships among multiple objectives. Most problems come when one grades essays. If the teacher establishes an "ideal essay" and grades against that standard the results will not be as realistic as if the teacher grades holistically. Remember that an essay should inform us about what the student understands and how he or she uses that understanding and, further, we can not compare a student level of understanding to that of a teacher. On a 10 point scale an outstanding essay may be worth 10, excellent ,9; good,8; fair, 7; poor, 6; and failing somewhat less, but to give 0-3 on an essay where there has been an honest attempt and some minimal display of knowledge is not realistic. Finally, in grading essays you may certainly correct spelling and grammar, but you grade content only-- a test does not provide opportunity for using dictionaries or for drafting and re-writing essays.

Published tests are not necessarily good. Any published test reflects the objectives of the author and not the teacher. A test must reflect the objectives which have been taught by the teacher, not printed in the book. This warning extends to the use of computer banks as well. I once saw a Biology teacher ask the computer to generate a 25 multiple choice item test on a unit, selecting random items. The completed test did not contain a single item on something the teacher spent two full days on in class!

Test design is an important issue for any teacher, supervisor, or administrator. Rieck's Rules on the matter are:

1) Each written test should contain both objective and essay items. This includes mathematics.

2) Every objective that is taught must be assessed on the test, providing the objective can be assessed on a test.

3) Nothing should be included on a test which has not been taught by the teacher.

4) The test should contain 75% of the items which are single objective items at the knowledge or comprehension level and 25% of the items which are on the higher order levels of Bloom's Taxonomy.

Both objective and essay items are included to provide students with different types of opportunities to demonstrate competence. In addition the use of essay items enhances the probability of sufficiently testing higher order skills.

If something is important enough to teach, it is important enough to assess. To omit test items which relate to objectives is to admit that the item was

not important. If this occurs too frequently students will ignore much of what you do in class because they know they may not be held responsible for it.

If something was not important enough to teach, it is not important enough to test. Just remember how much you hated people who tested you on footnotes or text content which was never part of the class discussion.

A test is designed to evaluate progress and diagnose difficulties. A test is not a game to see how "tough" we can be or how high a failure rate we can get. A test which allows the typical student to earn a "C" while providing a sufficient number of more difficult items to differentiate between A and B, and B and C is the best way to go.

It is never appropriate to curve. A well designed test which is criterion referenced is all that is necessary. Curing implies that you have an "n" of 100 or more who took the test under the same conditions. It also assumes that there must be a certain percentage of A's, B's, etc. This is nonsense! Students get what they earn, and if all earn A's great! On the other hand, if all earn F's that is also what they get, but we had better re-teach the material.

Portfolio and Other Assessments

Portfolio assessment simply implies a collection of student work samples to determine if objective have been achieved. High School administrators in New Jersey are very familiar with this since they must essentially do a portfolio assessment for students who fail to pass the state proficiency test prior to graduation. A committee, by using a wide variety of other acceptable measures, can certify that the student has achieved the required objectives even though the student did not pass the test.

Portfolios have been used by art teachers for some time. English teachers now collect samples of work. Science teachers can collect lab reports, projects, etc. By examining many forms of student work and looking for achievement it is possible to ascertain that a student has achieved minimum competence. Placing value on the samples, along with tests, etc can result in the assignment of a specific grade.

Pop quizzes, book reports, and countless other devices can be part of a total formal assessment program. These are too numerous to discuss in detail, but it is wise to remember that research shows frequent and multiple assessment enhances the learning process.

Informal Assessment

Teachers need to remember that body language is a great way to assess if people are learning. So is observing students when working at their seats or in cooperative learning groups. Teachers should never be stationary or behind

a desk during guided practice or group work. Teachers must be actively involved in monitoring performance with an eye toward re-teaching or adjusting instruction when difficulty is found.

Frequent use of questions in class is an excellent informal assessment. Such questions should frequently be open ended in nature. Student responses should be greeted warmly and never dismissed. Incorrect responses may be "shaped" toward correct ones; perhaps ques should be given to help students respond. To maximize the use of questioning it is important to remember "state the question--wait 5 seconds--identify a student." By asking the question without identifying a student the entire class will try to think of a response. If the student is identified first, the rest of the class is free to "tune out" knowing that, at least for the time being, they are "off the hook."

Homework is a time for error, it is not a time for grading. A major purpose of regular homework assignments is to help the student gain practice and re-enforce what was learned in class. The next day the homework must be used in some way as a tool to informally evaluate total class progress, and to diagnose where there may be difficulties which the teacher can address.

Role of Administration and Supervision

Because it may be true that many teachers are not good assessors, it becomes important to help teachers grow. Examine teachers' tests and make sure:

1) All objectives in the plan book are covered.

2) Nothing is on the test which does not appear in the plans.

3) The test has a reasonable design.

In the event a teacher's test does not meet the desired criteria, or if the total assessment plan is not satisfactory, it is appropriate to make specific suggestions for growth. In a management by objective format (such as the Professional Improvement Plan used in New Jersey) it is easy to:

1) Establish a growth goal or objective.

2) Indicate why there is a problem, stating specific reasons.

3) Establish specific steps necessary for the desired growth.

4) Establish time lines for the accomplishment of each step.

5) Establish criteria to determine success.

References

Andrews, Arthur (1983). "Grade Inflation-How Great? What are the Concerns of Parents, Educators." NASSP Bulletin 67,466.

Astin, Alexander W, et al (1979). "The American Freshman: National Norms for Fall 1979." Washington: American Council on Education.

Bangert-Drowns, R.L., Kulik, J.A., and Kulik, C.L. (1986). "Effects of Frequent Classroom Testing." Paper presented at the annual convention of the American Educational Research Association, San Francisco.

Cangelosi, James S. (1992). Systematic Teaching Strategies. New York: Longman.

Clark, Leonard H. and Starr, Irving S. (1991) . Secondary and Middle School Teaching Methods. New York: Macmillan Publishing Company.

Creager, John A. et al. (1969). "National Norms for Entering College Freshman-Fall 1969." Washington: American Council on Education.

Dillon-Peterson, Betty (1986). "Trusting Teachers to Know What's Good for Them." in Improving Teaching - 1986 ASCD Yearbook edited by Karen K. Zumwalt. Alexandria, VA; Association for Supervision and Curriculum Development.

Ellett, Chad D., Loup, Karen S. and Chauvin, Sheila W. (1990). "Annotated Guide to Teaching and Learning." Baton Rouge, LA: Louisiana Department of Education.

Fuchs, Lynn, Deno, Stanley L. and Mirkim, Phyllis (1984). "The Effects of Frequent Curriculum Based Measurement and Evaluation on Pedagogy, Student Achievement, and Student Learning." American Educational Research Journal 21,2.

Gronlund, N.E. (1973). Preparing Criterion Referenced for Classroom Use. New York: Macmillan.

Kindsvatter, Richard, Wilen, William, and Ishler, Margaret (1992). Dynamics of Effective Teaching. New York: Longman.

Manatt, Richard (1988). "Teacher Performance Evaluation: A Total Systems Approach." In Teacher Evaluation: Six Prescriptions for Success. Edited by Sarah J. Stanley and W. James Popham. Alexandria, VA.: Association for Supervision and Curriculum Development.

Mintz, S.L. (1979). "Teacher Planning: A Simulation Study." ERIC ED170276.

Oosterhof, Albert (1990). Classroom Applications of Educational Measurement. Columbus, Ohio: Merrill Publishing Company.

Peterson, P.L., Marx, R.W., and Clark, C.M. (1978). "Teacher Planning, Teacher Behavior, and Student's Achievement." American Educational Research Journal 15,3.

Popham, James W. and Hamberton, Ronald K. (1990). "Can You Pass the Test on Testing?" Principal 69,3.

Rieck, William A. (1989). "PIP = Instructional Improvement." Educational Viewpoints 9,1.

_____, (1989, October). "Teachers' Tests: An Aspect of Evaluation." NASSP Bulletin .

_____(1991) . "A Study of the Pedagogical Preparation in Undergraduate Pre-Service Programs for Secondary Education Majors in the United States." Lafayette, LA: University of Southwestern Louisiana.

_____(1992). "Student Assessment: How Supervisors Can Help (1992, August). ERIC Clearning House EA 023 978.

Sanders, James R. (1989). "Joint Committee for Standards for Teacher Competence in Educational Assessment of Students." Educational Measurement.

Schafer, William D. and Lissitz, Robert W. (1987). "Measurement Training for School Personnel: Recommendations and Reality." Journal of Teacher Education 38,3.

Stiggins, Richard J (1988). "Revitalizing Classroom Assessment." Phi Delta Kappan 69,5.

Terwilliger, James S. (1989). "Classroom Standard Setting and Grading Practices." Educational Measurement: Issues and Practice 8,2.